Econometric Theory, **11**, 1995, 736–749. Printed in the United States of America.

SPURIOUS BREAK

LUIS C. NUNES
University of Illinois
and
Universidade Nova de Lisboa

CHUNG-MING KUAN
University of Illinois
and
National Taiwan University

PAUL NEWBOLD
University of Illinois
and
University of Nottingham

A quasi-maximum likelihood estimator of the break date is analyzed. Consistency of the estimator is demonstrated under very general conditions, provided that the data-generating process is not integrated. However, the asymptotic distribution of the estimator is quite different for time series that are integrated of order one. In that case, when there is no break, the analyst can be spuriously led to the estimation of a break near the middle of the time series.

1. INTRODUCTION

The importance of considering structural change in statistical models is well documented in the literature (see the surveys by Zacks, 1983, and Krishnaiah and Miao, 1988). The problem is to test whether or not a change in the parameters of the model has occurred and, if so, to estimate when and by how much. The problem of estimation has been studied for several models with different estimation techniques (see Hinkley, 1970; Yao, 1987; among others). Also, several tests have been proposed treating the break date as unknown (see Brown, Durbin, and Evans, 1975, and Hawkins, 1987, among others, for independent and identically distributed [i.i.d.] observations or simple regression models, and Andrews, 1993, and Chu and White, 1992, among others, for more general models). However, Chu and White (1992)

We thank two referees and the co-editor for very helpful comments. Research was supported in part by a scholarship from Programa Ciência-JNICT Portugal. Address correspondence to: Chung-Ming Kuan, Department of Economics, 21 Hsu-Chow Road, National Taiwan University, Taipei 10020, Taiwan.

  0266-4666/95 $9.00 + .10

gave examples of tests for changing trend such that when a series is generated by an integrated process of order one ($I(1)$) with drift, but without structural change, the null hypothesis of no structural change will be rejected far too often.

A similar problem occurs when testing for unit roots in economic time-series where it is necessary to know in advance whether there has been a structural change or not. Perron (1989) showed that standard tests of the unit root hypothesis against trend-stationary alternatives reject the unit root hypothesis too infrequently if the true data-generating process (DGP) is that of stationary fluctuations around a trend function that contains a one-time break.

These dual problems occur because a time-series generated by an $I(1)$ process is very difficult to distinguish from one generated by a stationary process ($I(0)$) with structural change. This is well illustrated in Hendry and Neale (1991).

In this paper we concentrate on the estimation of the break date. We show that the quasi-maximum likelihood estimator (QMLE) for the break date is consistent under very weak assumptions on the error process provided it is *not* integrated of positive order. This extends previous theorems that require independence and normality of the errors (e.g., Krishnaiah and Miao, 1988). We also allow for trending regressors.

There is no study of the effect of integrated processes on the distribution of the estimator of the break date. In this paper we derive the asymptotic distribution of the break date QMLE in this case. We show that when a variable is generated by an $I(1)$ process without any structural change, the estimation of a model with structural change will suggest a spurious break. This result is analogous to the spurious regression and spurious trend results in Granger and Newbold (1974), Phillips (1986), and Durlauf and Phillips (1988) and complements the findings of Chu and White (1992).

In Section 2 we introduce a model of structural change and the QMLE of the break date. Asymptotic distributional results for the case where the error term is not integrated are given in Section 3. Section 4 derives results for the case where the error term is integrated of order one. In Section 5 we report simulation evidence confirming that our theoretical results provide useful insights into what would be found in the analysis of samples of moderate size. Proofs of the main theorems are given in the Appendix.

2. STRUCTURAL CHANGE MODEL

Given observations $\{y_t, x_t'\}$, where x_t is $p \times 1$, suppose we model the DGP of y_t as

$$y_t = x_t'\beta_t + \epsilon_t, \tag{1}$$

where $\{\epsilon_t\}$ is an $I(0)$ stochastic process. Consider a general change function for β_t:

$$\beta_t = \beta + g^{(T)}(t/T), \qquad t = 1,2,\dots,T, \tag{2}$$

where $g^{(T)}(\lambda)$ is a function of $\lambda \in [0,1]$ and may depend on the sample size T. With the prior belief that there is a single structural change occurring at some unknown point k_0, the model to be estimated can be written as

$$y_t = \begin{cases} x_t'\beta_1 + \epsilon_t, & t = 1,\dots,k_0, \\ x_t'\beta_2 + \epsilon_t, & t = k_0 + 1,\dots,T. \end{cases} \tag{3}$$

This corresponds to $\beta = \beta_1$ and $g^{(T)}(\lambda) = (\beta_2 - \beta_1)\mathbf{1}_{\{\lambda > \lambda_0\}}$ in (2), where $\mathbf{1}$ denotes the indicator function and $k_0 = [T\lambda_0]$ the integer part of $T\lambda_0$. The quasi-log-likelihood function is given by

$$-\frac{T}{2}\ln 2\pi - \frac{T}{2}\ln \sigma_\epsilon^2 - \frac{1}{2\sigma_\epsilon^2}\left(\sum_{t=1}^{k}(y_t - x_t'\beta_1)^2 + \sum_{t=k+1}^{T}(y_t - x_t'\beta_2)^2\right).$$

Conditional on k, the QMLE of β_1, β_2, and σ_ϵ^2 are

$$\hat\beta_1(k) = \left(\sum_{t=1}^{k} x_t x_t'\right)^{-1}\left(\sum_{t=1}^{k} x_t y_t\right),$$

$$\hat\beta_2(k) = \left(\sum_{t=k+1}^{T} x_t x_t'\right)^{-1}\left(\sum_{t=k+1}^{T} x_t y_t\right),$$

and $\hat\sigma_\epsilon^2(k) = \mathrm{RSS}_T(k)/T$, where

$$\mathrm{RSS}_T(k) = \left(\sum_{t=1}^{k}(y_t - x_t'\hat\beta_1(k))^2 + \sum_{t=k+1}^{T}(y_t - x_t'\hat\beta_2(k))^2\right). \tag{4}$$

The concentrated quasi-log-likelihood function is thus

$$-\frac{T}{2}\ln 2\pi - \frac{T}{2} - \frac{T}{2}\ln \hat\sigma_\epsilon^2(k).$$

The QMLE of k_0 is the integer $\hat k_T$, which solves

$$\min_{p \le k \le T-p} \hat\sigma_\epsilon^2(k) \quad \text{or} \quad \min_{p \le k \le T-p} \mathrm{RSS}_T(k). \tag{5}$$

Also define

$$\hat\lambda_T = \min\{\lambda : \lambda = \operatorname{argmin}_{u \in [\underline{\lambda},\bar\lambda]} \mathrm{RSS}_T([Tu])\}, \tag{6}$$

where $\underline{\lambda} < \bar\lambda$ are some prespecified constants in [0,1]. A reason for a choice of $[\underline{\lambda},\bar\lambda]$ other than the full interval [0,1] is that such a choice might yield significant precision gains if the change point is in, or close to, $[\underline{\lambda},\bar\lambda]$. Monte Carlo results in James, James, and Siegmund (1987) and Talwar (1983) for testing structural change in the location model suggest such a result. A more complete discussion of this issue in the context of testing for structural change can be found in Andrews (1993).

The following condition will be assumed to hold in the rest of the paper. Let D_T be a $p \times p$ diagonal matrix with diagonal elements d_{iT}, where d_{iT} are (possibly different) powers of T and $d_{iT}^{-1} \to 0$. This matrix is needed to normalize x_t properly because the components of x_t may be of different orders of magnitude in t. The condition is given by the following:

[A1] $D_T^{-1/2} \sum_{t=1}^{[T\lambda]} x_t x_t' D_T^{-1/2} \xrightarrow{p} Q(\lambda)$ uniformly in $\lambda \in [0,1]$ for some D_T, where $Q(\lambda)$ is positive-definite, symmetric, and an absolutely continuous, monotonically increasing function of λ; that is, $Q(\lambda_2) - Q(\lambda_1)$ is positive-definite for all $\lambda_2 > \lambda_1$.

Condition [A1] is a law of large numbers–type of condition. It excludes $I(1)$ processes. It can be seen that [A1] holds in the following leading examples.

Examples

1. Change in trend: $x_t = (1,t)'$, [A1] holds with

$$D_T = \begin{bmatrix} T & 0 \\ 0 & T^3 \end{bmatrix}, \qquad Q(\lambda) = \begin{bmatrix} \lambda & \lambda^2/2 \\ \lambda^2/2 & \lambda^3/3 \end{bmatrix}.$$

2. Change in regression coefficients: x_t is a $p \times 1$ random vector. Under standard regularity conditions,

$$\frac{1}{T} \sum_{t=1}^{[T\lambda]} x_t x_t' \xrightarrow{p} \lambda Q$$

uniformly in $\lambda \in [0,1]$, with Q a positive-definite matrix. It follows that [A1] holds with $D_T = TI_p$ and $Q(\lambda) = \lambda Q$.

3. CONSISTENCY

Under the assumption that the ϵ_t are independent and normally distributed, it has already been shown that the QMLE $\hat{\lambda}_T$ is consistent when there is a single structural change at some unknown $\lambda = \lambda_0 \in (0,1)$ (e.g., Krishnaiah and Miao, 1988). However, this result is based on too restrictive assumptions. Under more general conditions, we show that the break date estimator $\hat{\lambda}_T$ converges in probability to the true break date λ_0 when there is a single structural change at some $\lambda = \lambda_0 \in (0,1)$ or to the set $\{0,1\}$ if there is no structural change.

Consider the following condition:

[A2] $g^{(T)}(\lambda)$ is a function of bounded variation on $\lambda \in [0,1]$, which may depend on the sample size T, such that for D_T in [A1],

$$T^{-b/2} D_T^{1/2} g^{(T)}(\lambda) \to g^*(\lambda)$$

uniformly in $\lambda \in [0,1]$ for some function g^* on [0,1] and some constant $b > 0$.

Condition [A2] characterizes g^* as the limiting behavior of $g^{(T)}$; in particular, if $g^{(T)}(\lambda) = g(\lambda)$, then $g^*(\lambda) = F_p g(\lambda)$, where F_p is a selection matrix. It can be seen that [A2] holds in the following examples where a single change has occurred at some date $k_0 = [T\lambda_0]$ and $\lambda_0 \in (0,1)$ is a constant.

Examples

1. Change in trend:
   (a) Broken intercept: $y_t = \beta_1 + \delta_1 \mathbf{1}_{\{t \geq k_0+1\}} + \beta_2 t + \epsilon_t$. Since $g^{(T)}(\lambda) = (\delta_1, 0)' \mathbf{1}_{\{\lambda > \lambda_0\}} = g(\lambda)$, [A2] holds with $b = 1$ and $g^*(\lambda) = g(\lambda)$.
   (b) Broken discontinuous trend: $y_t = \beta_1 + \delta_1 \mathbf{1}_{\{t \geq k_0+1\}} + \beta_2 t + \delta_2 \mathbf{1}_{\{t \geq k_0+1\}} t + \epsilon_t$. Here, $g^{(T)}(\lambda) = (\delta_1, \delta_2)' \mathbf{1}_{\{\lambda > \lambda_0\}}$, and [A2] holds with $b = 3$ and

   $$g^*(\lambda) = (0, \delta_2)' \mathbf{1}_{\{\lambda > \lambda_0\}}.$$

   (c) Broken but continuous trend: $y_t = \beta_1 + \beta_2 t + \delta_2 \mathbf{1}_{\{t \geq k_0+1\}}(t - k_0) + \epsilon_t$. In this case, $g^{(T)}(\lambda) = (-\delta_2 [T\lambda_0], \delta_2)' \mathbf{1}_{\{\lambda > \lambda_0\}}$, and [A2] holds with $b = 3$ and

   $$g^*(\lambda) = (-\delta_2 \lambda_0, \delta_2)' \mathbf{1}_{\{\lambda > \lambda_0\}}.$$

2. Change in regression coefficients: Since $g^{(T)}(\lambda) = \delta \mathbf{1}_{\{\lambda > \lambda_0\}} = g(\lambda)$, [A2] holds with $b = 1$ and $g^*(\lambda) = g(\lambda)$.

Consider now the following condition:

[A3] For D_T in [A1],

$$\left(D_T^{-1/2} \sum_{t=1}^{[T\lambda]} x_t \epsilon_t,\ 0 \leq \lambda \leq 1 \right) \Rightarrow R,$$

where R is a p-dimensional Gaussian process on [0,1] with $R(0) = 0$ and having mean 0 and covariance $E[R(\lambda_1)R(\lambda_2)'] = \Sigma(\min(\lambda_1, \lambda_2))$, with

$$\Sigma(\lambda) = \lim_{T\to\infty} \frac{1}{T} E\left[\left(\sum_{t=1}^{[T\lambda]} x_t \epsilon_t \right) \left(\sum_{t=1}^{[T\lambda]} x_t \epsilon_t \right)' \right]$$

absolutely continuous for all λ.

This condition is general enough to allow ϵ_t and x_t to be weakly dependent, heterogeneous random variables but *not* integrated of positive order (see, e.g., Wooldridge and White, 1988).

Examples

1. Change in trend: Suppose that ϵ_t satisfies the regularity conditions in Phillips and Perron (1988) such that $\sigma^2 = \lim_T E(S_T^2/T) > 0$, where $S_T = \sum_{t=1}^T \epsilon_t$. Then,

   $$R(\lambda) = \sigma \int_0^\lambda (1, r)'\, dW(r),$$

   with W the standard one-dimensional Brownian motion. The variance of $R(\lambda)$ is

   $$\Sigma(\lambda) = \sigma^2 \int_0^\lambda (1, r)'(1, r)\, dr = \sigma^2 \int_0^\lambda \begin{bmatrix} 1 & r \\ r & r^2 \end{bmatrix} dr = \sigma^2 Q(\lambda).$$

2. Change in regression coefficients: Suppose now that ϵ_t and x_t satisfy the assumptions in Ploberger, Krämer, and Kontrus (1989). Then,

$$R(\lambda) = \sigma Q^{1/2} W(\lambda), \qquad \Sigma(\lambda) = \sigma^2 \lambda Q,$$

with W the standard p-dimensional Brownian motion.

THEOREM 3.1. *Given DGP* (1) *and* (2), *suppose that* [A1] *and* [A3] *are satisfied.*

1. *If* [A2] *holds with* $g^*(\lambda) = \delta \mathbf{1}_{\{\lambda > \lambda_0\}}$ *and* $\delta \in \mathbb{R}^p \setminus \{0\}$, *and* $0 < \underline{\lambda} < \lambda_0 < \bar{\lambda} < 1$, *then* $\hat{\lambda}_T \xrightarrow{p} \lambda_0$ *as* $T \to \infty$.
2. *Suppose there is no change, i.e.,* $g^{(T)}(\lambda) \equiv 0$.
   (a) *If* $0 < \underline{\lambda} < \bar{\lambda} < 1$, *then*

$$\hat{\lambda}_T \Rightarrow \mathrm{argmax}_{\lambda \in [\underline{\lambda}, \bar{\lambda}]} R(\lambda)' Q(\lambda)^{-1} R(\lambda) \\ + [R(1) - R(\lambda)]'[Q(1) - Q(\lambda)]^{-1}[R(1) - R(\lambda)].$$

   (b) *If* $\underline{\lambda} = 0$ *and* $\bar{\lambda} = 1$, *then* $\hat{\lambda}_T \xrightarrow{p} \{0,1\}$ *as* $T \to \infty$.

Remark 1. In part 1, δ can depend on λ_0. It is possible that part 1 could be relaxed to allow $\underline{\lambda} = 0$ and $\bar{\lambda} = 1$ but the proof of this would require different arguments, as in Hinkley (1970) and Yao (1987).

Remark 2. Part 1 extends previous results such as those of Hinkley (1970) and Krishnaiah and Miao (1988) in that x_t and ϵ_t may be weakly dependent and heterogeneously distributed; our result also allows for trending regressors. Comparing to Bai (1994), which concentrates on the mean change of linear processes, our condition [A3] has broader applicability, but we do not have a rate of convergence result.

Remark 3. Part 2(a) shows that when there is no structural change, consistency of $\hat{\lambda}_T$ is not possible if it is restricted as in (6); instead, $\hat{\lambda}_T$ has an asymptotic distribution with support equal to $[\underline{\lambda}, \bar{\lambda}]$ (see also Remark 1 after Theorem 4.1). We note that these results in fact formalize a discussion in Andrews (1993, p. 839).

4. SPURIOUS BREAK: y_t IS INTEGRATED OF ORDER ONE

Suppose now that the true DGP is characterized by no structural change, but ϵ_t is $I(1)$, violating condition [A3]. We show that when estimating a model with structural change the QMLE estimator $\hat{k}_T$ will suggest a spurious break. This is shown in Theorem 4.1 as well as in the simulations in Section 5.

Suppose then that we estimate model (3) as in Section 2 and that [A1] continues to hold for x_t. However, y_t is generated by an $I(1)$ process and there is no structural change. Instead of (1) and (2), the DGP for y_t is given by $y_0 = 0$ and

$$y_t = y_{t-1} + \eta_t = \sum_{j=1}^{t} \eta_j, \tag{7}$$

with $\{\eta_t\}$ and $I(0)$ stochastic process such that the following condition holds:

[A3′] For D_T in [A1] and some constant $\alpha > 0$,

$$\left(T^{-\alpha/2} D_T^{-1/2} \sum_{t=1}^{[T\lambda]} x_t y_t,\ 0 \le \lambda \le 1\right) \Rightarrow G,$$

where G is some p-dimensional functional of a Gaussian process.

The main difference between [A3] and [A3′] is the rate of convergence in the term $T^{-\alpha/2}$.

Examples

1. Change in trend: [A3′] holds with $\alpha = 2$ and

$$\begin{pmatrix} T^{-3/2} \sum_{t=1}^{[T\lambda]} y_t \\ T^{-5/2} \sum_{t=1}^{[T\lambda]} t y_t \end{pmatrix} \Rightarrow G(\lambda) := \sigma \int_0^{\lambda} \begin{pmatrix} 1 \\ r \end{pmatrix} W(r)\, dr.$$

2. Regression: [A3′] holds with $\alpha = 1$. Following Park and Phillips (1988), let $w_t = (x_t', \eta_t)'$ satisfy a multivariate invariance principle: $T^{-1/2} \sum_{j=1}^{[T\lambda]} w_j \Rightarrow B(\lambda)$, a $(p+1)$-vector Brownian motion $B = (B_1', B_2)'$ with covariance matrix $\Omega := V + \Lambda + \Lambda'$. Then,

$$T^{-1} \sum_{t=1}^{[T\lambda]} x_t y_t \Rightarrow G(\lambda) := \int_0^{\lambda} B_2\, dB_1 + \lambda \Delta_{21}',$$

where $\Delta_{21} = V_{21} + \Lambda_{21}$ is the corresponding lower left submatrix.

THEOREM 4.1. *Suppose the DGP for y_t is given by* (7), [A1] *and* [A3′] *are satisfied, and* $0 < \underline{\lambda} < \lambda_0 < \bar{\lambda} < 1$. *Then,*

$$\hat{\lambda}_T \Rightarrow \operatorname{argmax}_{\lambda \in [\underline{\lambda}, \bar{\lambda}]} G(\lambda)' Q(\lambda)^{-1} G(\lambda) + [G(1) - G(\lambda)]'[Q(1) - Q(\lambda)]^{-1}[G(1) - G(\lambda)].$$

Remark 1. The theorem says that $\hat{\lambda}_T$ has an asymptotic distribution with support equal to $[\underline{\lambda}, \bar{\lambda}]$; we were, however, unable to extend this result to the case where $\underline{\lambda} = 0$ and $\bar{\lambda} = 1$. Our simulation results (to be discussed in Section 5) show that, in contrast with the distribution of part 2 of Theorem 3.1, which is more concentrated in the tails, this distribution is more concentrated in the middle of the sample even when the support is enlarged to $[0,1]$. That is, it is more likely that $\hat{\lambda}_T$ would suggest a spurious break near the middle of the sample when y_t is $I(1)$ without a break, but $\hat{\lambda}_T$ is more likely to be close to $\underline{\lambda}$ and $\bar{\lambda}$ (or 0 and 1) when y_t is $I(0)$ without a break. This result is new but analogous to the well-known spurious regression and spurious trend results.

Remark 2. Suppose that $x_t = (x_{1t},\dots,x_{pt})'$. Condition [A3′] can be written as

$$\left(T^{-\alpha_i/2} d_{iT}^{-1/2} \sum_{t=1}^{[T\lambda]} x_{it} y_t,\ 0 \le \lambda \le 1\right) \Rightarrow G_i, \qquad i = 1,\dots,p,$$

with $\alpha_i = \alpha > 0$. If, however, the α_i are not all the same, then let $\alpha_* = \max_{1\le i\le p} \alpha_i$. Then Theorem 4.1 still holds with the elements of $G(1)$ and $G(\lambda)$ that correspond to $\alpha_i < \alpha_*$ replaced by zeros.

5. SIMULATIONS

In this section we discuss some simulations. The model used in the simulations is the change in mean model with autocorrelated disturbances:

$$y_t = \beta_t + \epsilon_t \quad \text{with } \beta_t = \delta \mathbf{1}_{\{t \ge [T\lambda]+1\}},$$

where $\epsilon_t = \phi\epsilon_{t-1} + \eta_t$ and η_t i.i.d. $N(0,1)$. The number of replications for each experiment was 100,000, and we have used the normal pseudo-random number generator in GAUSS-386 for i.i.d. $N(0,1)$ innovations. The results are presented in Figures 1–3. The break-point estimators are obtained by considering all possible break points as in (5) except in Figure 2b. We do not report our further results for the change in trend models nor for the model employed by Ploberger and Krämer (1992) and Andrews (1993) because all the results were qualitatively identical.

Figure 1 considers a change in the middle of the sample with $\delta = 1$, $\lambda = 0.5$, and $\phi = 0$. For $T = 100$, the estimator is concentrated around the true break date. Although not shown here, as T increases, the estimator becomes more precise. Also, when some autocorrelation is present (e.g., $\phi = 0.5$), the estimator loses precision when compared with the case of no autocorrelation, $\phi = 0$. We have also considered a change closer to the beginning of the sample (e.g., $\lambda = 0.25$). The same comments made for the previous case apply here. The only difference is that the mode of the distribution is shifted to the new break date, but with little change in the dispersion of the distribution.

For the case of no change with $I(0)$ errors, Figure 2a gives the distribution of $\hat{k}$ when it takes all possible values, whereas Figure 2b is the distribution when $\hat{k}$ is restricted as in (6) with $\underline{\lambda} = 0.15$ and $\bar{\lambda} = 0.85$. For $\phi = 0$, the mass of the distribution is more concentrated in the tails than in the middle. Also, although not shown here, when ϕ increases or T decreases, the mass tends to become less concentrated in the tails. The same conclusions hold when $\hat{k}$ is restricted for various $\underline{\lambda}$ and $\bar{\lambda}$.

Figure 3a illustrates the case of spurious break: no change but with $I(1)$ errors ($\delta = 0$, $\phi = 1$) for $T = 100$. As opposed to the second case, and more like the first case, the mass of the distribution is more concentrated in the middle than in the tails. The contrast between the graphs clearly suggests

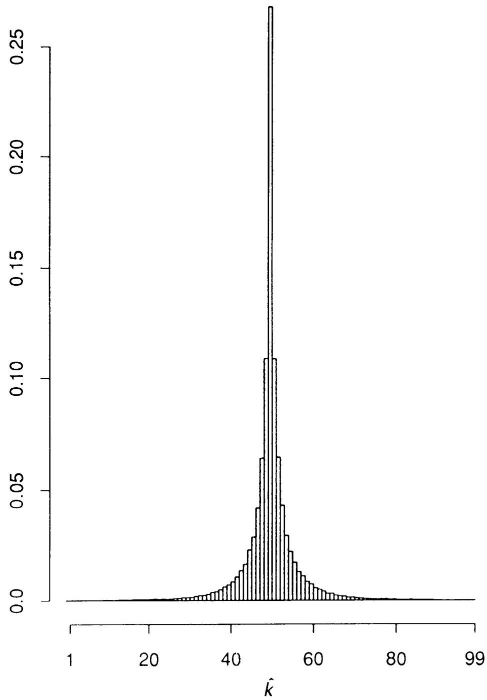

FIGURE 1. Distribution of $\hat{k}$ when $\delta = 1$, $\lambda = 0.5$, and $\phi = 0$ for $T = 100$.

what we mean by a spurious break. We have also simulated the asymptotic distribution, $T = \infty$, for the previous case. The density estimate was approximated using $T = 1{,}000$ and 100,000 replications, and the resulting empirical density function was smoothed with a Gaussian kernel with a bandwidth equal to 0.2. The result is shown in Figure 3b. Simulations also showed that the asymptotic distribution seems to be a very good approximation for values of T as low as 25.

In conclusion, the simulation results support the asymptotic results of Sections 3 and 4. Moreover, they suggest that the insights provided by the asymptotics are relevant for quite small sample sizes.

6. CONCLUSION

In this paper we have shown that the quasi-maximum likelihood break date estimator is consistent when the error term satisfies a condition that allows for general error processes *not* integrated of positive order. We also obtained the asymptotic distribution of the break date estimator when the series is generated by an $I(1)$ process. In this case, estimation of a structural change model will result in the appearance of a spurious break.

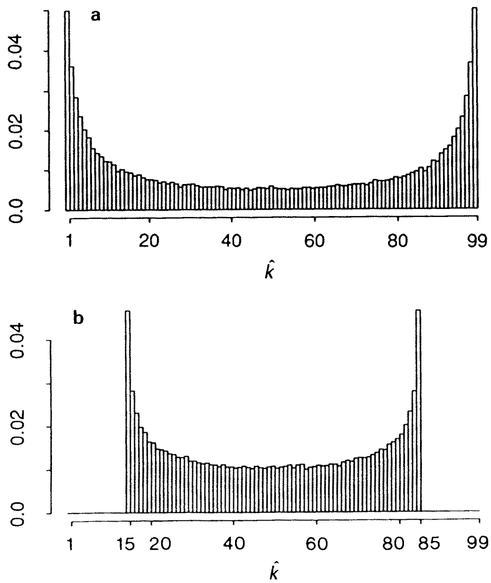

FIGURE 2. a. Distribution of $\hat{k}$ when $\delta = 0$ and $\phi = 0$ for $T = 100$. b. Distribution of $\hat{k}$ when $\delta = 0$ and $\phi = 0$ for $T = 100$ with $\underline{\lambda} = 0.15$ and $\bar{\lambda} = 0.85$.

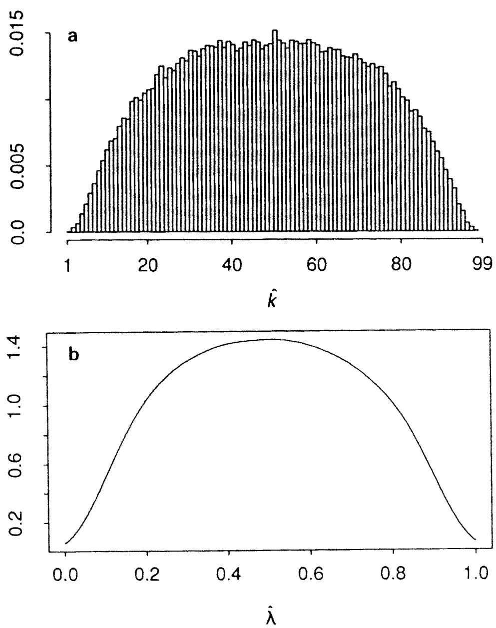

FIGURE 3. a. Distribution of $\hat{k}$ when $\delta = 0$ and $\phi = 1$ for $T = 100$. b. Distribution of $\hat{\lambda}$ when $\delta = 0$ and $\phi = 1$ for $T = 100$.

Our results, together with some of those cited in the introduction, suggest that disentangling information in data about both unit roots and structural change is likely to be a difficult task. This is essentially the graphical insight suggested by Hendry and Neale (1991). The important question suggested by such results is that of construction of tests for structural change and estimators of break dates that are valid irrespective of whether a unit root is present or not in the noise component. For an attempt in this direction, see Perron (1991).

REFERENCES

Andrews, D.W.K. (1993) Tests for parameter instability and structural change with unknown change point. *Econometrica* 61, 821–856.

Arnold, L. (1974) *Stochastic Differential Equations: Theory and Applications.* New York: Wiley.

Bai, J. (1994) Least squares estimation of a shift in linear processes. *Journal of Time Series Analysis* 15, 453–472.

Brown, R.L., J. Durbin, & J.M. Evans (1975) Techniques for testing the constancy of regression relationships over time. *Journal of the Royal Statistical Society, Series B* 37, 149–163.

Chu, C.-S.J. & H. White (1992) A direct test for changing trend. *Journal of Business and Economic Statistics* 10, 289–299.

Durlauf, S.N. & P.C.B. Phillips (1988) Trends versus random walks in time series analysis. *Econometrica* 56, 1333–1354.

Granger, C.W.J. & P. Newbold (1974) Spurious regressions in econometrics. *Journal of Econometrics* 2, 111–120.

Hawkins, D.L. (1987) A test for a change point in a parametric model based on a maximal Wald-type statistic. *Sankhyā* 49, 368–376.

Hendry, D.F. & A.J. Neale (1991) A Monte Carlo study of the effects of structural breaks on tests for unit roots. In P. Hackl & A.H. Westlund (eds.), *Economic Structural Change: Analysis and Forecasting*, pp. 95–119. New York: Springer-Verlag.

Hinkley, D. (1970) Inference about the change point in a sequence of random variables. *Biometrika* 57, 1–17.

James, B., K.L. James, & D. Siegmund (1987) Tests for a change-point. *Biometrika* 74, 71–83.

Krämer, W., W. Ploberger, & R. Alt (1988) Testing for structural change in dynamic models. *Econometrica* 56, 1355–1369.

Krishnaiah, P.R. & B.Q. Miao (1988) Review about estimation of change points. In P.R. Krishnaiah & C.R. Rao (eds.), *Handbook of Statistics*, vol. 7, pp. 375–402. New York: Elsevier.

Park, J.Y. & P.C.B. Phillips (1988) Statistical inference in regressions with integrated processes: Part 1. *Econometric Theory* 4, 468–497.

Perron, P. (1989) The great crash, the oil price shock and the unit root hypothesis. *Econometrica* 57, 1361–1401.

Perron, P. (1991) A Test for Changes in a Polynomial Trend Function for a Dynamic Time Series. Manuscript, Princeton University.

Phillips, P.C.B. (1986) Understanding spurious regressions in econometrics. *Journal of Econometrics* 33, 311–340.

Phillips, P.C.B. & P. Perron (1988) Testing for a unit root in time series regression. *Biometrika* 75, 335–346.

Ploberger, W. & W. Krämer (1992) The CUSUM test with OLS residuals. *Econometrica* 60, 271–285.

Ploberger, W., W. Krämer, & K. Kontrus (1989) A new test for structural stability in the linear regression model. *Journal of Econometrics* 40, 307–318.

Talwar, P.P. (1983) Detecting a shift in location. *Journal of Econometrics* 23, 353–367.

Wooldridge, J.M. & H. White (1988) Some invariance principles and central limit theorems for dependent heterogeneous processes. *Econometric Theory* 4, 210–230.
Yao, Y.-C. (1987) Approximating the distribution of the ML estimate of the change-point in a sequence of independent r.v.'s. *Annals of Statistics* 15, 1321–1328.
Zacks, S. (1983) Survey of classical and Bayesian approaches to the change point problem: Fixed sample and sequential procedures for testing and estimation. In M.H. Rivzi et al. (eds.), *Recent Advances in Statistics*, pp. 245–269. New York: Academic Press.

APPENDIX

LEMMA A.1. *Given* [A1] *and* [A2], *we have*

$$T^{-b/2}D_T^{-1/2}\sum_{t=1}^{[T\lambda]} x_t x_t' g^{(T)}\left(\frac{t}{T}\right) \xrightarrow{p} \int_0^\lambda dQ(r)g^*(r), \tag{8}$$

$$T^{-b/2}D_T^{-1/2}\sum_{t=[T\lambda]+1}^{T} x_t x_t' g^{(T)}\left(\frac{t}{T}\right) \xrightarrow{p} \int_\lambda^1 dQ(r)g^*(r), \tag{9}$$

as $T \to \infty$, *uniformly in* $\lambda \in [0,1]$.

Proof. From [A1], [A2], and equation (28) of Krämer, Ploberger, and Alt (1988),

$$T^{-b/2}D_T^{-1/2}\sum_{t=1}^{[T\lambda]} x_t x_t' g^{(T)}\left(\frac{t}{T}\right) = \sum_{t=1}^{[T\lambda]} D_T^{-1/2}x_t x_t' D_T^{-1/2}T^{-b/2}D_T^{1/2}g^{(T)}\left(\frac{t}{T}\right)$$
$$\xrightarrow{p} \int_0^\lambda dQ(r)g^*(r).$$

This proves (8). The proof of (9) is similar. ■

Proof of Part 1 of Theorem 3.1. Let

$$M_T(k) = \left(\sum_{t=1}^{k} \epsilon_t^* x_t'\right)\left(\sum_{t=1}^{k} x_t x_t'\right)^{-1}\left(\sum_{t=1}^{k} x_t \epsilon_t^*\right) + \left(\sum_{t=k+1}^{T} \epsilon_t^* x_t'\right)\left(\sum_{t=k+1}^{T} x_t x_t'\right)^{-1}\left(\sum_{t=k+1}^{T} x_t \epsilon_t^*\right) \tag{10}$$

and

$$\epsilon_t^* = \epsilon_t + x_t' g^{(T)}\left(\frac{t}{T}\right). \tag{11}$$

It is easy to see that $\mathrm{RSS}_T(k) = \sum_{t=1}^T \epsilon_t^{*2} - M_T(k)$ so that

$$\hat\lambda_T = \min\{\lambda : \lambda = \mathrm{argmax}_{u\in[\underline{\lambda},\bar\lambda]} M_T([Tu])\}. \tag{12}$$

Using [A1]–[A3], it follows after some manipulations and Lemma A.1 that

$$T^{-b}M_T([T\lambda]) \xrightarrow{p} L(\lambda). \tag{13}$$

as $T \to \infty$, uniformly in $\lambda \in [\underline{\lambda}, \bar{\lambda}]$, where

$$L(\lambda) = \left[\int_0^\lambda dQ(r)g^*(r)\right]' Q(\lambda)^{-1}\left[\int_0^\lambda dQ(r)g^*(r)\right]$$
$$+ \left[\int_\lambda^1 dQ(r)g^*(r)\right]' [Q(1) - Q(\lambda)]^{-1}\left[\int_\lambda^1 dQ(r)g^*(r)\right].$$

Using the definition of $g^*(\lambda)$,

$$L(\lambda) = \begin{cases} [(Q(1) - Q(\lambda_0))\delta]'[Q(1) - Q(\lambda)]^{-1}[(Q(1) - Q(\lambda_0))\delta], & \lambda \le \lambda_0, \\ [(Q(\lambda) - Q(\lambda_0))\delta]' Q(\lambda)^{-1}[(Q(\lambda) - Q(\lambda_0))\delta] & \\ \quad + [(Q(1) - Q(\lambda))\delta]'[Q(1) - Q(\lambda)]^{-1}[(Q(1) - Q(\lambda))\delta], & \lambda > \lambda_0, \end{cases}$$
$$= \begin{cases} \delta'([Q(1) - Q(\lambda_0)][Q(1) - Q(\lambda)]^{-1}[Q(1) - Q(\lambda_0)])\delta, & \lambda \le \lambda_0, \\ \delta'(Q(1) - 2Q(\lambda_0) + Q(\lambda_0)Q(\lambda)^{-1}Q(\lambda_0))\delta, & \lambda > \lambda_0. \end{cases}$$

Because by [A1] $Q(\lambda_2) - Q(\lambda_1)$ is positive-definite for all $\lambda_2 > \lambda_1$, it is easy to see that a maximum of $L(\lambda)$ occurs exactly at $\lambda = \lambda_0$ for any $\delta \neq 0$. Because the convergence in (13) is uniform in $\lambda \in [\underline{\lambda}, \bar{\lambda}]$, this proves the theorem. ∎

Proof of Part 2 of Theorem 3.1. Because there is no change, $\epsilon_t^* = \epsilon_t$ in (11). Thus, by [A1], [A3], and the continuous mapping theorem, we have for $\underline{\lambda} \le \lambda \le \bar{\lambda}$

$$M_T([T\lambda]) \Rightarrow M(\lambda), \tag{14}$$

where $M(\lambda) \equiv R(\lambda)'Q(\lambda)^{-1}R(\lambda) + [R(1) - R(\lambda)]'[Q(1) - Q(\lambda)]^{-1}[R(1) - R(\lambda)]$. Another application of the continuous mapping theorem gives subpart (a). By [A1] $q(\lambda) := dQ(\lambda)/d\lambda$ is a nonnegative-definite symmetric matrix. By L'Hospital's rule,

$$\lim_{\lambda \to 0} \frac{Q(\lambda)}{\lambda} = q(0). \tag{15}$$

Also, by [A3], let $d\Sigma(\lambda)/d\lambda = P(\lambda)P(\lambda)'$ with $P(\lambda) = U(\lambda)\Lambda(\lambda)^{1/2}U(\lambda)'$, where $\Lambda(\lambda)$ is the diagonal matrix of eigenvalues and $U(\lambda)$ the orthogonal matrix of column eigenvectors of $d\Sigma(\lambda)/d\lambda$. Then the process

$$Z(\lambda) = \int_0^\lambda P(r)\, dW(r),$$

where W is the standard p-dimensional Brownian motion, has the same distribution as $R(\lambda)$. With probability one (w.p.1), the set of limit points of the net

$$\left\{ \frac{Z(\lambda)}{\sqrt{2\lambda \log\log\left(\frac{1}{\lambda}\right)}} \right\}_{e^{-1} > \lambda \downarrow 0}$$

is equal to the closed ellipsoid

$$E_0 = P(0)H_p = \{x \in \mathbb{R}^p : x = P(0)y,\ y \in H_p\},$$

where $H_p = \{x \in \mathbb{R}^p : |x| \leq 1\}$ (see Chapter 7 of Arnold, 1974). This result together with (15) implies that w.p.1

$$\limsup_{\lambda \to 0} \frac{R(\lambda)' Q(\lambda)^{-1} R(\lambda)}{2 \log \log \left(\frac{1}{\lambda}\right)}$$

$$= \limsup_{\lambda \to 0} \frac{R(\lambda)'}{\sqrt{2\lambda \log \log \left(\frac{1}{\lambda}\right)}} \left(\frac{Q(\lambda)}{\lambda}\right)^{-1} \frac{R(\lambda)}{\sqrt{2\lambda \log \log \left(\frac{1}{\lambda}\right)}}$$

$$= \max_i \gamma_i,$$

where γ_i are the eigenvalues of $P(0) q(0)^{-1} P(0)'$, provided that $q(0)$ is invertible. Then, w.p.1

$$\limsup_{\lambda \to 0} R(\lambda)' Q(\lambda)^{-1} R(\lambda) = \infty. \tag{16}$$

Similarly, it can be shown that w.p.1

$$\limsup_{\lambda \to 1} [R(1) - R(\lambda)]' [Q(1) - Q(\lambda)]^{-1} [R(1) - R(\lambda)] = \infty. \tag{17}$$

If $q(0)$ is not invertible, the same conclusions hold by considering the Moore–Penrose inverse of $q(0)$. Following an argument similar to the proof of Corollary 1 of Andrews (1993), it follows that $\sup_{\lambda \in [0,1]} M_T([T\lambda]) \xrightarrow{p} \infty$. But, because $\sup_{\lambda \in [\underline{\lambda}, \bar{\lambda}]} M_T([T\lambda]) = O_p(1)$ for $0 < \underline{\lambda} < \bar{\lambda} < 1$, then $\hat{\lambda}_T \xrightarrow{p} \{0,1\}$ as $T \to \infty$ when $\underline{\lambda} = 0$ and $\bar{\lambda} = 1$. This proves subpart (b). ∎

Proof of Theorem 4.1. From (7) it is easy to see that (12) and (10) hold with $\epsilon_t^* = y_t$ in (10). In (12), substitute $M_T([Tu])$ by $T^{-\alpha} M_T([Tu])$. This does not change $\hat{\lambda}_T$ for any T because α does not depend on u. From (10), we have

$$T^{-\alpha} M_T(k)$$

$$= \left(\sum_{t=1}^{k} y_t x_t' D_T^{-1/2} T^{-\alpha/2}\right) \left(D_T^{-1/2} \sum_{t=1}^{k} x_t x_t' D_T^{-1/2}\right)^{-1} \left(T^{-\alpha/2} D_T^{-1/2} \sum_{t=1}^{k} x_t y_t\right)$$

$$+ \left(\sum_{t=k+1}^{T} y_t x_t' D_T^{-1/2} T^{-\alpha/2}\right) \left(D_T^{-1/2} \sum_{t=k+1}^{k} x_t x_t' D_T^{-1/2}\right)^{-1} \left(T^{-\alpha/2} D_T^{-1/2} \sum_{t=k+1}^{T} x_t y_t\right).$$

By [A1], [A3′], and the continuous mapping theorem, we have for $\underline{\lambda} \leq \lambda \leq \bar{\lambda}$

$$T^{-\alpha} M_T([T\lambda]) \Rightarrow G(\lambda)' Q(\lambda)^{-1} G(\lambda)$$
$$+ [G(1) - G(\lambda)]' [Q(1) - Q(\lambda)]^{-1} [G(1) - G(\lambda)].$$

The proof is completed by another application of the continuous mapping theorem. ∎

LINKED CITATIONS

Spurious Break

Luis C. Nunes; Chung-Ming Kuan; Paul Newbold

Econometric Theory, Vol. 11, No. 4. (Oct., 1995), pp. 736-749.

Stable URL:

http://links.jstor.org/sici?sici=0266-4666%28199510%2911%3A4%3C736%3ASB%3E2.0.CO%3B2-Z

References

Tests for Parameter Instability and Structural Change With Unknown Change Point

Donald W. K. Andrews

Econometrica, Vol. 61, No. 4. (Jul., 1993), pp. 821-856.

Stable URL:

http://links.jstor.org/sici?sici=0012-9682%28199307%2961%3A4%3C821%3ATFPIAS%3E2.0.CO%3B2-I

Techniques for Testing the Constancy of Regression Relationships over Time

R. L. Brown; J. Durbin; J. M. Evans

Journal of the Royal Statistical Society. Series B (Methodological), Vol. 37, No. 2. (1975), pp. 149-192.

Stable URL:

http://links.jstor.org/sici?sici=0035-9246%281975%2937%3A2%3C149%3ATFTTCO%3E2.0.CO%3B2-7

A Direct Test for Changing Trend

Chia-Shang James Chu; Halbert White

Journal of Business & Economic Statistics, Vol. 10, No. 3. (Jul., 1992), pp. 289-299.

Stable URL:

http://links.jstor.org/sici?sici=0735-0015%28199207%2910%3A3%3C289%3AADTFCT%3E2.0.CO%3B2-P

Trends versus Random Walks in Time Series Analysis

Steven N. Durlauf; Peter C. B. Phillips

Econometrica, Vol. 56, No. 6. (Nov., 1988), pp. 1333-1354.

Stable URL:

http://links.jstor.org/sici?sici=0012-9682%28198811%2956%3A6%3C1333%3ATVRWIT%3E2.0.CO%3B2-5

LINKED CITATIONS

Inference About the Change-Point in a Sequence of Random Variables
David V. Hinkley
Biometrika, Vol. 57, No. 1. (Apr., 1970), pp. 1-17.
Stable URL:
http://links.jstor.org/sici?sici=0006-3444%28197004%2957%3A1%3C1%3AIATCIA%3E2.0.CO%3B2-9

Tests for a Change-Point
Barry James; Kang Ling James; David Siegmund
Biometrika, Vol. 74, No. 1. (Mar., 1987), pp. 71-83.
Stable URL:
http://links.jstor.org/sici?sici=0006-3444%28198703%2974%3A1%3C71%3ATFAC%3E2.0.CO%3B2-2

Testing for Structural Change in Dynamic Models
Walter Krämer; Werner Ploberger; Raimund Alt
Econometrica, Vol. 56, No. 6. (Nov., 1988), pp. 1355-1369.
Stable URL:
http://links.jstor.org/sici?sici=0012-9682%28198811%2956%3A6%3C1355%3ATFSCID%3E2.0.CO%3B2-4

Statistical Inference in Regressions with Integrated Processes: Part 1
Joon Y. Park; Peter C. B. Phillips
Econometric Theory, Vol. 4, No. 3. (Dec., 1988), pp. 468-497.
Stable URL:
http://links.jstor.org/sici?sici=0266-4666%28198812%294%3A3%3C468%3ASIIRWI%3E2.0.CO%3B2-D

The Great Crash, the Oil Price Shock, and the Unit Root Hypothesis
Pierre Perron
Econometrica, Vol. 57, No. 6. (Nov., 1989), pp. 1361-1401.
Stable URL:
http://links.jstor.org/sici?sici=0012-9682%28198911%2957%3A6%3C1361%3ATGCTOP%3E2.0.CO%3B2-W

Testing for a Unit Root in Time Series Regression
Peter C. B. Phillips; Pierre Perron
Biometrika, Vol. 75, No. 2. (Jun., 1988), pp. 335-346.
Stable URL:
http://links.jstor.org/sici?sici=0006-3444%28198806%2975%3A2%3C335%3ATFAURI%3E2.0.CO%3B2-B

LINKED CITATIONS

The Cusum Test with Ols Residuals
Werner Ploberger; Walter Krämer
Econometrica, Vol. 60, No. 2. (Mar., 1992), pp. 271-285.
Stable URL:
http://links.jstor.org/sici?sici=0012-9682%28199203%2960%3A2%3C271%3ATCTWOR%3E2.0.CO%3B2-G

Some Invariance Principles and Central Limit Theorems for Dependent Heterogeneous Processes
Jeffrey M. Wooldridge; Halbert White
Econometric Theory, Vol. 4, No. 2. (Aug., 1988), pp. 210-230.
Stable URL:
http://links.jstor.org/sici?sici=0266-4666%28198808%294%3A2%3C210%3ASIPACL%3E2.0.CO%3B2-H

Approximating the Distribution of the Maximum Likelihood Estimate of the Change-Point in a Sequence of Independent Random Variables
Yi-Ching Yao
The Annals of Statistics, Vol. 15, No. 3. (Sep., 1987), pp. 1321-1328.
Stable URL:
http://links.jstor.org/sici?sici=0090-5364%28198709%2915%3A3%3C1321%3AATDOTM%3E2.0.CO%3B2-8